



МГУ имени Н.Э. Баумана

Кафедра ИУ-1 «Системы автоматического управления»

Методы вычислений

Численные методы поиска экстремумов функций



Андрей Леонидович Масленников
amas@bmstu.ru

2023 г.

Задача оптимизации

Метод оптимизации

Метод поиска экстремума функции

Задача оптимизации — это задача при решении которой требуется добиться минимума (реже максимума) некоторого критерия при заданных ограничениях.

Метод оптимизации — это, фактически, математическая постановка задачи оптимизации в которой задается критерий J , теоретически показываются и обосновываются, доказываются свойства сходимости, точности и др.

Численный метод поиска экстремума функции — это вычислительный алгоритм, который позволяет подобрать такой \mathbf{x} , при котором достигается J_{\min}

Векторная форма постановки задачи

решение \longrightarrow $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{X}} \|\mathbf{J}(\mathbf{x})\| = \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_k J_k(\mathbf{x})$ критерий

область допустимых значений аргумента \longleftarrow

по типу поиска \mathbf{x} :

- локальные;
- глобальные.

по характеру задачи оптимизации:

- детерминированные;
- стохастические;
- смешанные.

по гладкости $f(\mathbf{x})$ и существованию производных:

- прямые;
- первого порядка — требуется знание $\nabla f(\mathbf{x})$;
- второго порядка — требуется знание $\nabla^2 f(\mathbf{x})$.

одномерные:

- метод полного перебора;
- метод дихотомии;
- метод золотого сечения;
- и другие.



прямые:

- метод Гаусса;
- метод Хука—Дживса;
- метод Нелдера—Мида;
- и другие.

первого порядка:

- метод покоординатного спуска;
- метод градиентного спуска;
- метод наискорейшего спуска.

второго порядка:

- метод Ньютона;
- метод Ньютона—Рафсона;
- квази-Ньютоновские методы;
- метод Левенберга—Марквардта;
- и другие.

Метод полного перебора — это метод поиска экстремума функции, в котором в интервале поиска $[a, b]$ формируется n равноотстоящих точек значений аргумента x , для которых последовательно вычисляются значения функции $f(x)$, после чего из полученного набора выбирается минимальное значение $f_{\min}(x_{\min})$ и соответствующее значение аргумента x_{\min} .

Точность поиска экстремума составляет

$$\varepsilon = \frac{1}{2} \frac{b-a}{n}$$

Метод крайне вычислительно
не эффективен

while $k < n$

$$x_k = a + k \frac{b-a}{n+1}$$

if $f(x_k) < f_{\min}(x)$

$$f_{\min}(x) = f(x_k)$$

$$x_{\min} = x_k$$

end

$$k = k + 1$$

end

Метод дихотомии — это метод поиска экстремума функции, в котором интервале поиска $[a, b]$ после каждой итерации уменьшается в два раза

Критерий остановки

$$|b - a| < \varepsilon$$

Значение δ

$$\delta \in \left(0, \frac{b - a}{2}\right)$$

while $|b - a| > \varepsilon$

$$x_k = \frac{a + b}{2}$$

if $f(x_k - \delta) \geq f(x_k + \delta)$

$$a = x_k$$

else

$$b = x_k$$

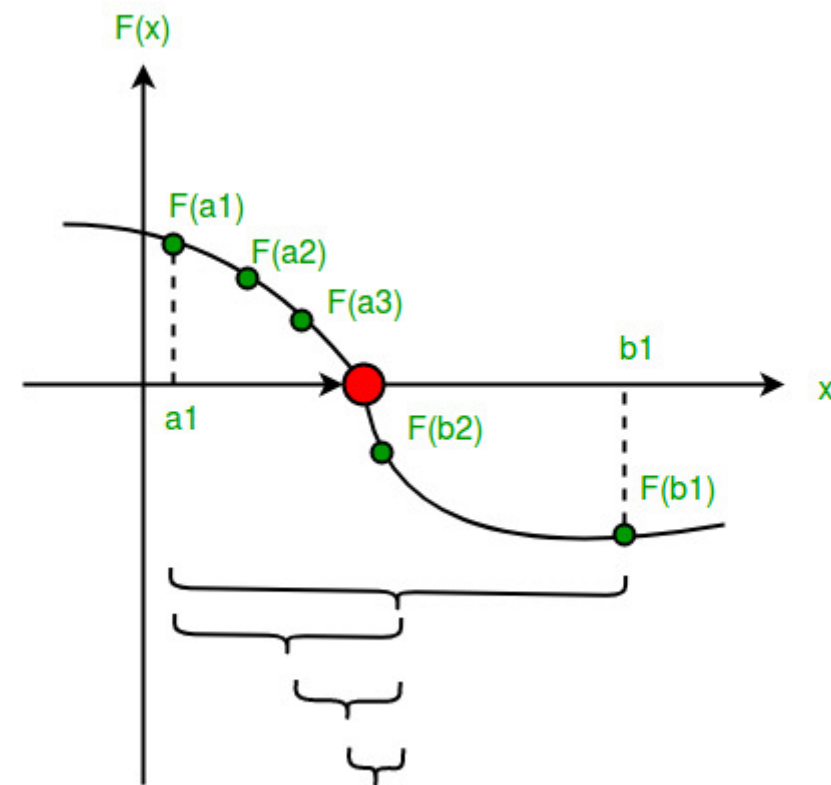
end

$$k = k + 1$$

end

$$x_{\min} = \frac{a + b}{2}$$

$$f_{\min} = f(x_{\min})$$



Начальные условия

$$f(x) = (x-1)^2$$

$$[a, b] = [0.0, 1.5]$$

$$\varepsilon = 0.1$$

$$\delta = \frac{b-a}{4}$$

Таблица решения

a	b	δ	x	$f(x - \delta)$		$f(x + \delta)$	$ b - a $
0	1.500	0.3750	0.7500	0.3906	>	0.0156	0.7500
0.7500	1.500	0.1875	1.1250	0.0039	<	0.0977	0.3750
0.7500	1.1250	0.0938	0.9375	0.0244	>	0.0010	0.1875
0.9375	1.1250	0.0469	1.0313	0.0002	<	0.0061	0.0938
0.9375	1.0313						

Итоговое решение

$$x_{\min} = \frac{0.9375 + 1.0313}{2} = 0.9844$$

$$f(x_{\min}) = 0.0002$$

меньше ε

Метод золотого сечения — это метод поиска экстремума функции, в котором интервал поиска $[a, b]$ после каждой итерации уменьшается пропорционально числу φ .

Критерий остановки

$$|b - a| < \varepsilon$$

Значение φ

$$\varphi = \frac{1 + \sqrt{5}}{2} \approx 1.62$$

$$x_1 = b - \frac{b - a}{\varphi}$$

$$x_2 = a + \frac{b - a}{\varphi}$$

while $|b - a| > \varepsilon$

if $f(x_1) \geq f(x_2)$

$$a = x_1, \quad x_1 = x_2, \quad x_2 = \dots$$

else

$$b = x_2, \quad x_2 = x_1, \quad x_1 = \dots$$

end

end

$$x_{\min} = \frac{a + b}{2}$$

$$f_{\min} = f(x_{\min})$$

Начальные условия

$$f(x) = (x-1)^2$$

$$[a, b] = [0.4, 1.5]$$

$$\varepsilon = 0.1$$

Таблица решения

a	b	x_1	x_2	$f(x_1)$		$f(x_2)$	$ b - a $
0.4000	1.5000	0.8202	1.0798	0.0323	>	0.0064	0.6798
0.8202	1.5000	1.0798	1.2403	0.0064	<	0.0578	0.4202
0.8202	1.2403	0.9807	1.0798	0.0004	>	0.0064	0.2597
0.8202	1.0798	0.9193	0.9807	0.0065	<	0.0004	0.1605
0.9193	1.0798	0.9807	1.0185	0.0004	>	0.0003	0.0992
0.9807	1.0798						

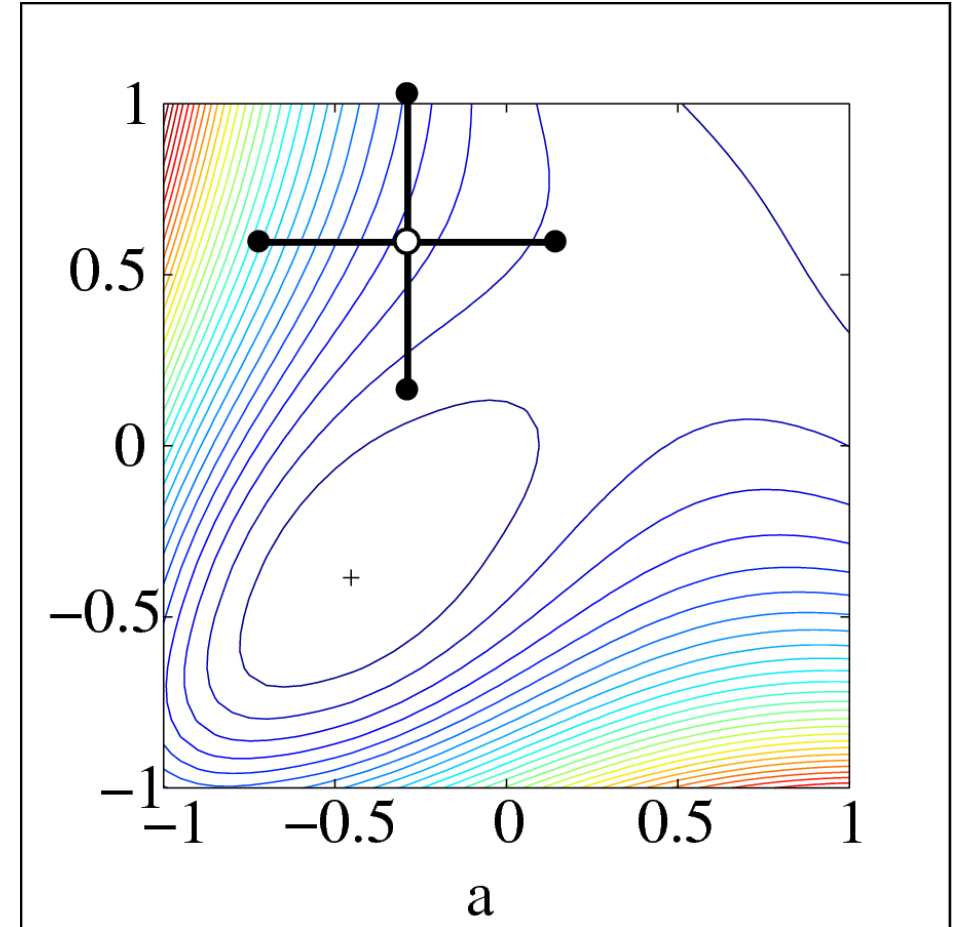
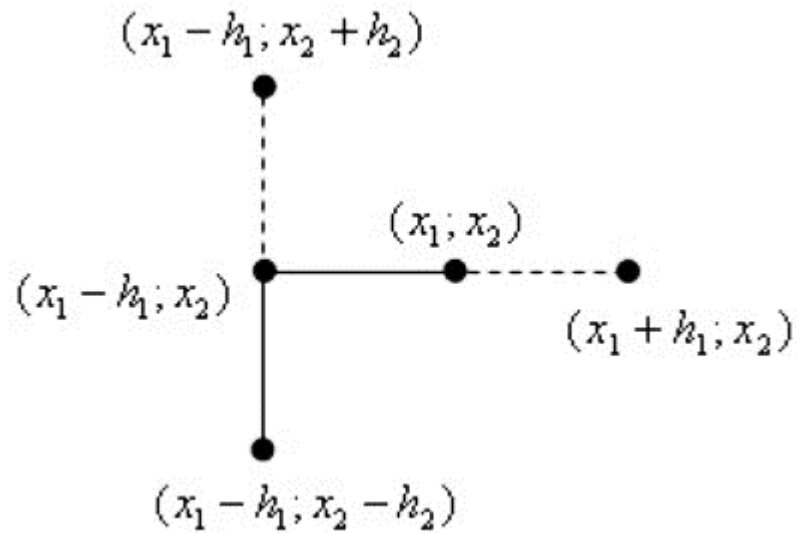
Итоговое решение

$$x_{\min} = \frac{0.9807 + 1.0798}{2} = 1.0302$$

$$f(x_{\min}) = 9.1469 \times 10^{-4}$$

меньше ε

Метод Хука—Дживса — это метод n -мерной сетке вычисляются значения функции в точках, отстающих от заданной базовой точки на величину λ , последовательно по каждой координате с перемещением в точку с минимальным значением функции. λ для каждого x_i постепенно уменьшается.

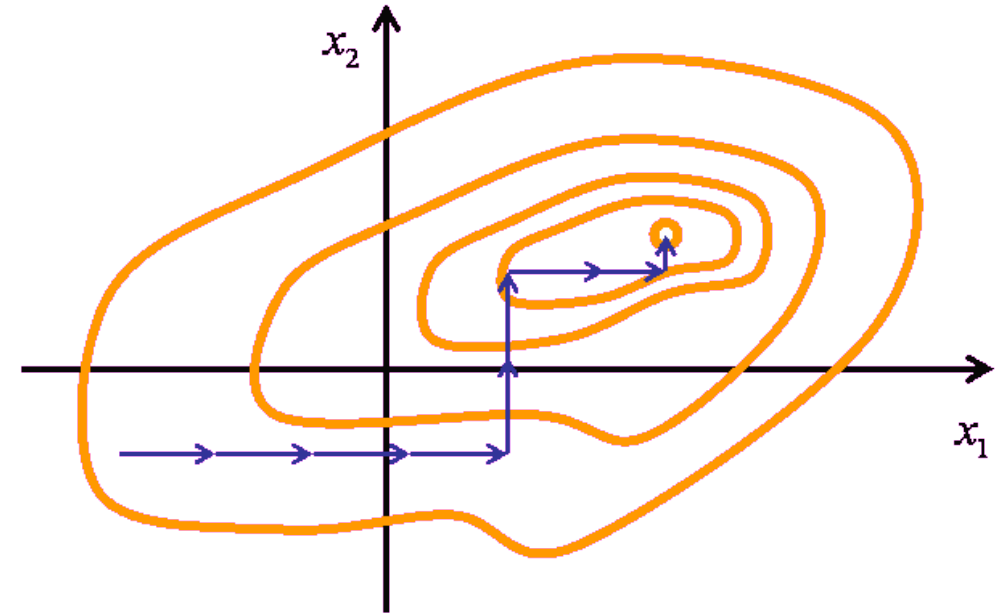


Формируется итерационный процесс на каждой k -ой итерации которого последовательно по каждой компоненте вектора x сдвигается базовая точка:

$$x_{i,k} = x_{i,k-1} + \lambda_{i,k} e_i$$

где e_i — формируют ортонормированный базис, а $\lambda_{i,k}$ находится решением своей оптимизационной задачи

$$\lambda_{i,k} = \arg \min_{\lambda_{i,k}} f(x_{i,k-1} + \lambda_{i,k} e_i)$$



Отличительная особенность метода состоит в том, что направление движения по каждой координате — фиксировано

Метод покоординатного спуска

$$x_{i,k} = x_{i,k-1} + \lambda_{i,k} \nabla f_i(\mathbf{x}_{k-1})$$

как в методе Гаусса, только направление поиска не постоянно, а определяется градиентом

Метод градиентного спуска

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \lambda_k \nabla f(\mathbf{x}_{k-1})$$

движение к минимуму идет сразу по всем координатам одновременно

где

λ_k — шаг перехода в сторону минимума;

$\nabla f(\mathbf{x}_k)$ — вектор градиентов функции, который определяется как

$$\nabla f(\mathbf{x}_k) = - \left. \frac{\partial f(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}_k}$$

только в этом случае градиентный метод называется **методом наискорейшего спуска**

Варианты задания λ

$$\lambda = \text{const}$$

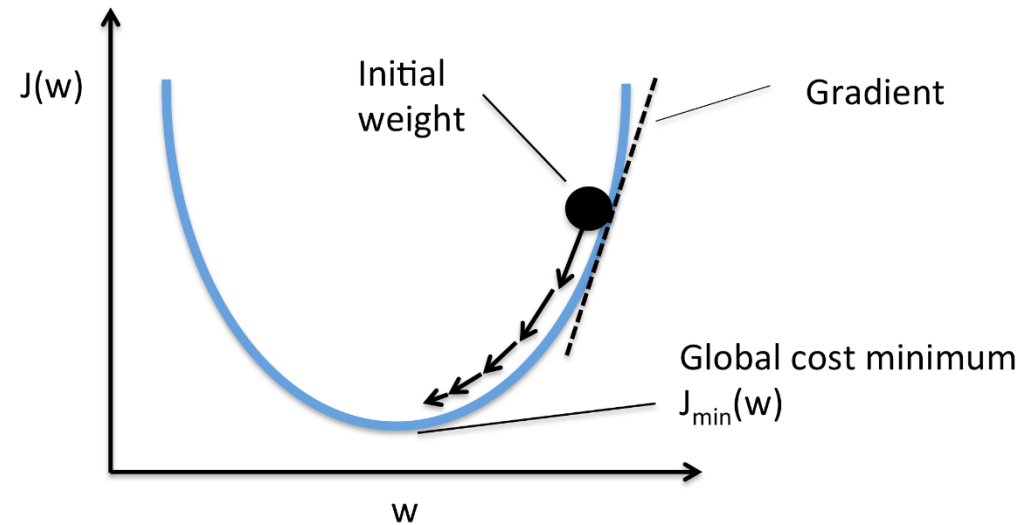
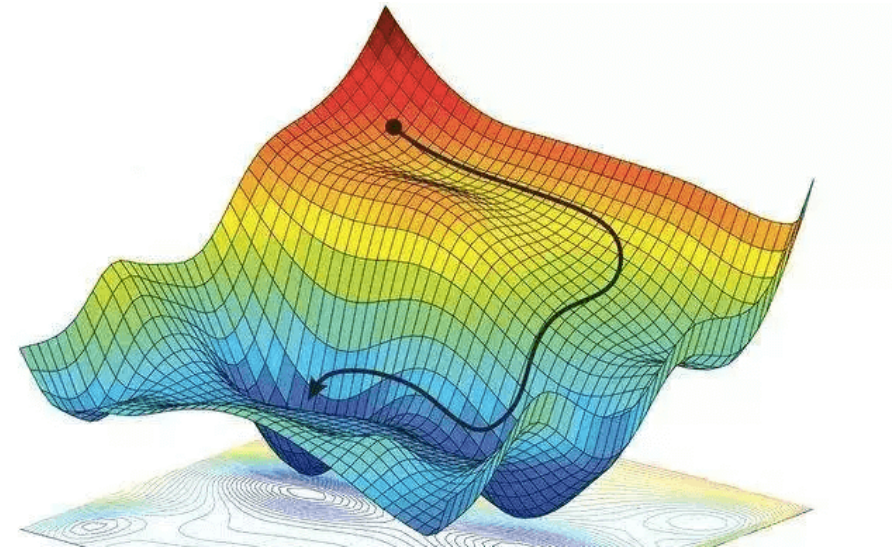
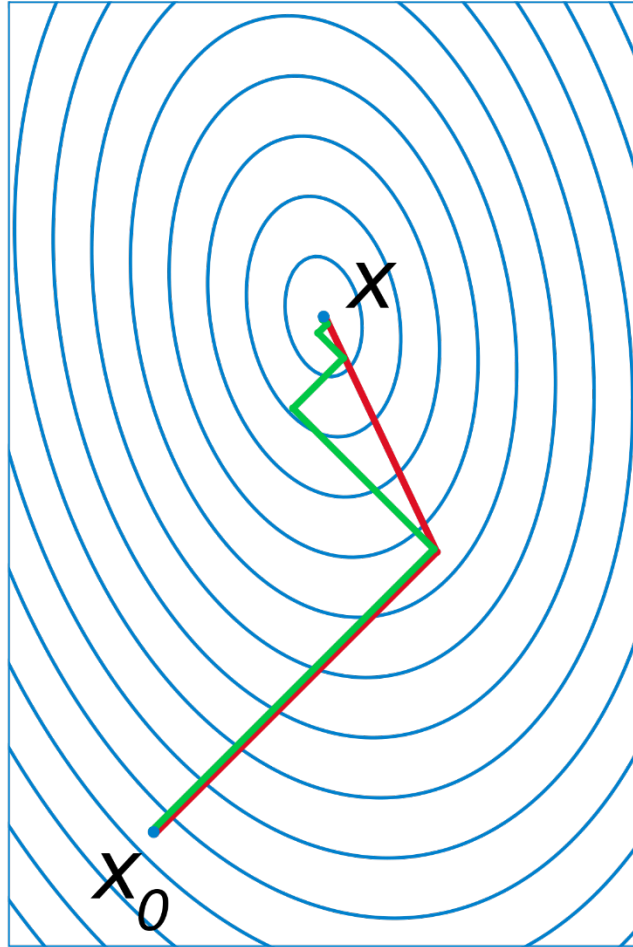
(как в методе Гаусса, но может расходиться)

$$\lambda_k = \frac{\lambda_{k-1}}{p}$$

(уменьшается с каждым шагом, следовательно может не сходиться к минимуму)

$$\lambda_k = \underset{\lambda_k}{\operatorname{argmin}} f(\mathbf{x}_{k-1} + \lambda_k \nabla f(\mathbf{x}_{k-1}))$$

(решается через скалярную задачу оптимизации)



Алгоритм покоординатных методов

```

while criteria ok
  for  $i = 1 : n$ 
     $\nabla f_i(\mathbf{x}_k) = \dots$ 
     $\lambda_{i,k} = \dots$ 
     $x_{i,k} = \dots$ 
  end
   $k = k + 1$ 
end

```

Алгоритм непокоординатных методов

```

while criteria ok
   $\nabla f(\mathbf{x}_k) = \dots$ 
   $\lambda_k = \dots$ 
   $\mathbf{x}_k = \dots$ 
   $k = k + 1$ 
end

```

Выход по значению функции

$$|f(\mathbf{x}_k) - f(\mathbf{x}_{k-1})| < \varepsilon_f$$

Выход по значению аргумента

$$\|\mathbf{x}_k - \mathbf{x}_{k-1}\| < \varepsilon_x$$

Выход по объему вычислений

$$k > k_{\max}$$

$$k_f > k_{f, \max}$$

максимальное количество
вычислений значений
функции $f(\mathbf{x}_k)$

Методы второго порядка — это метод поиска экстремума функции нескольких переменных, шаг поиска минимума в которых определяется матрицей Гессе

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{H}^{-1}(\mathbf{x}_{k-1}) \nabla f(\mathbf{x}_{k-1})$$

Метод Ньютона—Рафсона

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{H}^{-1}(\mathbf{x}_k) \nabla f(\mathbf{x}_k)$$

а λ_k находится решением оптимизационной задачи:

$$\lambda_k = \arg \min_{\lambda_k} f\left(\mathbf{x}_k + \lambda_k \mathbf{H}^{-1}(\mathbf{x}_k) \nabla f(\mathbf{x}_k)\right)$$

Матрица Гессе

$$\mathbf{H}(\mathbf{x}_k) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix}_{\mathbf{x}_k}$$

Проблемы применения методов второго порядка

- вычисление матрицы Гессе;
(большой объем вычислений, особенно если численно)
- вычисление обратной матрицы Гессе;
(большой объем вычислений, потенциальная неустойчивость)



Квази-Ньютоновские методы

Методы наименьших квадратов — это оптимизации, в котором минимизируется среднеквадратичная величина ошибки

$$\sum_i e^2 = \sum_i (f(\mathbf{x}_i) - y_i)^2 \rightarrow \min$$

Аппроксимация экспериментальных данных — это численный подбор параметров (вектор θ) функции $f(\theta, t)$, наиболее точно описывающих (повторяющих) характер изменения набора данных $y(t)$, полученного экспериментально.

$$\theta = \arg \min_{\theta \in \Theta} \sum_i (f(\theta, t_i) - y_i(t_i))^2$$

